Supporting Information

Knoch et al. 10.1073/pnas.0911619106

SI Materials and Methods

Subjects. We studied 87 right-handed men (mean age, 22.6 years; range, 20–27 years). All provided written informed consent to participate in the study, which was approved by the local Ethics Committee. None of the subjects had been previously subjected to TMS or a trust game. No subject had a history of psychiatric illness or neurologic disorder. There was no difference among the 6 experimental groups with respect to age ($\chi^2 = 8.478$; df = 5; P = .1318; Kruskal-Wallis test). None of the subjects experienced serious adverse side effects or reported scalp pain, neck pain, or headache after the experiment.

Transcranial Magnetic Stimulation. rTMS was administered to the DLPFC for 15 min before subjects participated in the trust game ("off-line paradigm") (1), using a Magstim Rapid Magnetic Stimulator and a commercially available figure-of-eight coil (70-mm-diameter double-circle, air-cooled). For stimulation of the right and left DLPFC, the TMS coil was placed over F4 and F3 using the electroencephalogram 10–20 coordination system, as in previous studies (2-4). We chose this approach because no previous fMRI data on this paradigm exist. We used the real-time neuronavigation option for BrainVoyager QX 1.6 with the Zebris CMS20S measuring system for real-time motion analysis (Zebris Medical GmbH) to ensure correct placement of the TMS coil during the stimulation. The stimulation intensity was set at 54% of the maximum stimulator output. The coil was held tangential to the subject's head with the handle pointing rostrally. Subjects received a single 15-min, 1-Hz rTMS train (900 pulses) over either the left DLPFC or right DLPFC, or sham stimulation using a Magstim placebo coil, which looks identical to the real coil and also delivers the characteristic "click" sound. Half of the subjects in the sham stimulation group received sham rTMS over the right DLPFC, and half received it over the left DLPFC.

The rTMS parameters are well within currently recommended guidelines (5) and result in suppression of excitability of the targeted cortical region for several minutes after completion of the rTMS train (1). Subjects performed the task immediately after the end of the stimulation train in the same laboratory room. Because the subjects received the instructions for the game before the rTMS train, it was possible to begin with the task about 30 s after completion of the stimulation train, thus under the influence of the rTMS aftereffect.

The Trust Game Experiment. Subjects received instructions that explained the rules of the game before stimulation. Each subject was required to complete a series of test questions after reading the instructions to verify comprehension. The experiment consisted of 15 identical periods of a trust game. Each period was divided into 2 stages, an investor decision stage and a trustee decision stage. In the first stage, the investor was endowed with 10 points and had to decide the amount that he wished to invest in the current trustee. His choice was restricted to 1, 4, 7, or 10 points. (Restricting the possible choices reduces the cognitive effort of evaluating the game situation and simplifies tabulating the payoffs.) Zero investments were not allowed, to exclude "costless" reputation formation. Also note that no exact medium investment was allowed; that is, investors were forced to choose either a high or a low investment. Once the decision was made, the invested amount was quadrupled and passed to the trustee (after pretesting, we decided to quadruple the amount—instead of the usual tripling—to ensure sufficient 10-point investments in the anonymous condition, allowing for profound statistical inference); that is, the received amount was 4-fold greater than the invested amount. In the second stage, the trustee then had to decide how much of the quadrupled amount he wanted to transfer back to the investor. This back-transfer was not quadrupled. The trustee's choice was restricted to the following:

(i) Back-transfer nothing (= 0% of the received amount)

(ii) Back-transfer 25% of the received amount (= the invested amount). The investor finishes the period with his endowment of 10 points, allowing him to "break even."

(iii) Transfer an amount that equalizes payoffs between the 2 subjects (= 62.5% of the received amount). The resulting payoff table is as presented in Table S2.

Restricting the trustee's choices to 3 options (which are all fixed percentages of the amount transferred by the investor) serves 2 purposes. First, the number of available options is independent of the actual investment, so cognitive effort is kept constant across investment levels. Second, it makes the information about choice available to the trustee uninformative about the investment level; that is, investors can be told the trustees' past choices without revealing the size of the investments. After each period, the trustees were randomly rematched with another investor. There were 2 information conditions: investors either had no information about the current trustee's past choices (anonymous condition) or were informed about the current trustee's 3 past choices on their decision screen (reputation condition). Investors could only observe choices (eg, "transfer nothing" or "equalize payoffs"); they had no information about the size of the corresponding investments or the chronological order of the choices. These 3 design features—a stable information window of 3 periods, no information on investment levels, and no information about the sequence of choicesensured that each of the trustee's decisions (given past decisions) had exactly the same reputational relevance, thus keeping the strategic incentive associated with the choice constant. Moreover, these features helped keep the cognitive effort of the game

The *immediate* benefit of the trustee's choice, however, was proportional to the amount that the investor transferred, as the payoff table illustrates. This made it possible to observe different levels of "temptation."

Subjects took part in only the reputation condition or the anonymity condition. Only trustees received stimulation. Because TMS could be applied to only one subject at a time, investors came collectively to the laboratory of the Institute for Empirical Research in Economics ("investor sessions"), while trustees were located at the University Hospital of Zurich ("trustee sessions"). Implementation was such that investors and additional trustees completed the experiment in the computer lab before the TMS experiment. The subjects receiving stimulation then came individually to the University Hospital Zurich for the trustee sessions. In each period, each of these subjects was randomly matched with a subject from the investor sessions such that their histories were matched; for example, a trustee who had opted twice for "nothing" and once for "equalize" in the 3 previous periods would be matched with an investor who had observed the same play in the 3 previous periods in the investor session. The average net duration of the experiment (from onset of the trust game to completion of control questions) was 401.5 s (maximum, 478.8 s). All participants were paid according to their payoffs in the game; 1 point in the game equaled 0.20 Swiss francs (CHF). In addition, participants in the lab received a show-up fee of CHF 10, and participants receiving stimulation received a show-up fee of CHF 60. We implemented the experiment in this way to ensure that subjects had a monetary incentive and a real concern about reputation.

Measuring Subjects' Fairness Judgments. Because we hypothesized that rTMS to the right DLPFC would foster unfair behavior, we were interested in exploring whether rTMS also has a similar impact on the judgment of unfair behavior or whether the notion of fairness remains unchanged, thus creating a gap between judgment and choice. Directly after the completion of the experiment, subjects had to answer the following question: "Please indicate in the following how you evaluate the participant's behavior: Assume that a participant A has transferred 7 points to a participant B. Participant B then chose the option 'transfer nothing.' How do you evaluate the fairness of participant B's behavior?" Responses to this question were given on a 7-point Likert scale ranging from 1 ("very unfair") to 7 ("very fair").

Measuring Subjects' Ability to Assess the Future Consequences of Past Back-Transfer Behaviors. One alternative explanation for our findings could be that rTMS of the right DLPFC does not remove the ability to override immediate short-run benefits, but simply affects subjects' ability to assess the future consequences of previous back-transfer behaviors. To examine this possibility, we asked subjects how many points (1, 4, 7, or 10) they would expect an investor to transfer to a trustee who had opted twice for "equalize payoff" and once for back-transferring nothing in the previous 3 periods. The exact wording of the question was as follows: "Assume that in the previous 3 periods, participant B has chosen once not to transfer anything, and has chosen twice to 'equalize payoffs.' How much do you expect the investor with whom B is matched next to transfer to B?" Subjects' response options were 1, 4, 7, or 10 points.

Fairness Judgment and Assessment of Future Consequences Across Treatments. As mentioned in the main article, we found no significant differences across stimulation conditions in the answers to the 2 questions posed earlier. The bar graphs in Figs. S1 and S2 illustrate the mean responses to the fairness judgment question and the hypothetical investment question by stimulation condition.

Measuring Subjects' Impulsivity and Personal Norm of Reciprocity. Approximately 10 days after the experiment, we sent those participants who were assigned the role of trustee a questionnaire that included the BIS and BAS scales developed by Carver and White (24 items) and translated into German by Strobel et al. (6, 7). This inventory investigates a subject's impulsive reaction to aversive stimuli (BIS) and rewarding stimuli (BAS). The questionnaire also included the 27-item Personal Norm of Reciprocity scale of Perugini et al. (8). Items in this inventory explore a subject's tendency to reward another person's positive behavior (positive reciprocity) and to punish negative behavior (negative reciprocity). This questionnaire was translated at our institute and checked by back-translation.

Statistical Analyses. To investigate whether the decision of how much to back-transfer differed across stimulation conditions, we used regression models, all estimated in STATA version 10 (StataCorp). In the regression models, the dependent variable is the fraction of received points that the trustee transfers back. The dummy variables "right DLPFC" and "left DLPFC" are included to model the baseline effect of the 3 stimulation conditions (condition "sham" is the omitted category in our specification; results do not change if we use "right DLPFC" or "left DLPFC" as the omitted category). As explained in the main

article, the temptation to defect, and thus the self-control effort required for reputation formation, is largest if the investor transfers 10 points to the trustee. Thus, we hypothesize that the required recruitment of right DLPFC is highest in this case, implying that disruption of right DLPFC function is more likely to generate a behavioral effect. For this reason, our regressions also include a dummy variable for all those observations in which the investor actually transferred 10 points. The differences in the effect of stimulation conditions on the back-transfer decision (for all situations in which the investor was sending 10 points) are then estimated by including the corresponding interaction terms: We interact each stimulation condition (right DLPFC, left DLPFC, and sham) with a dummy variable indicating an investment of 10 points. For example, the interaction term between right DLPFC and an investment of 10 (denoted "right DLPFC × investment = 10") then identifies the difference in the effect on back-transfers between the right DLPFC condition and the sham stimulation condition (ie, the omitted category) for the investment = 10 case. Note that the incentive for reputation formation disappears in the last period because the interaction between investors and trustees does not continue after period 15. Thus, we dropped the last observed period (period 15) from the data for all experimental conditions and for all observed experimen-

Tables S3 and S4 give the results of GLS estimation for the reputation condition and the anonymous condition, respectively. We present linear regression models in these tables, but the results obtained by the corresponding ordered discrete choice models for the 3 choice categories of the investor are identical with the results given here. We use robust standard errors adjusted for clustering on the subject level. We also include a random effect—assumed to be normally distributed—for each subject in our sample. We estimated various specifications of the regression model for each experimental condition (reputation and anonymous), and report 4 specifications here. Because the variable "sham DLPFC \times investment = 10" is the omitted category in specification 1, specification 1 tests whether the effect of stimulation of the right DLPFC and the left DLPFC differs from the effect of sham stimulation. In specification 2, the variable "left DLPFC \times investment = 10" is the omitted category; this specification tests whether the effect of stimulation of the right DLPFC and sham stimulation differs from the effect of stimulation of the left DLPFC. Specifications 3 and 4 differ from specifications 1 and 2 because we also controlled for individual subjects' fairness judgments, impulsivity, and reciprocity norm.

In Table S3, we see that the interaction term "right DLPFC \times investment = 10" is highly significantly negative in all 4 specifications. Specification 1 shows that, controlling for baseline differences captured by the variables right DLPFC and left DLPFC, in the investment = 10 case, subjects whose right DLPFC was stimulated back-transferred 11.4 percentage points less than subjects in the sham stimulation condition. Similarly, specification 2 shows that subjects in the right DLPFC condition back-transferred 19.8 percentage points less than subjects in the left DLPFC stimulation condition. This indicates that regardless of whether we compare the effect of stimulation of the right DLPFC with either the sham stimulation or stimulation of the left DLPFC, stimulation of the right DLPFC always had a significantly negative effect on the trustee's back-transfers. But the effects of left DLPFC and sham stimulation on backtransfers did not differ significantly from each other, as demonstrated by the insignificant coefficient estimates of the dummy variables "left DLPFC \times investment = 10" and "sham \times investment = 10."

As noted above, specifications 3 and 4 include further control variables and reveal that those who judged the scenario described earlier to more fair than others tended to back-transfer less.

Including the control variables does not mitigate the strength of the effect of right DLPFC stimulation on back-transfers, however.

Table S4 shows the same 4 specifications as in Table S3 for data from the anonymous condition. We see no significant differences in the effects of stimulation of the right DLPFC, stimulation of the left DLPFC, and sham stimulation on backtransfers, because the estimated coefficients of the variables "right DLPFC × investment = 10," "left DLPFC × Investment = 10," and "sham DLPFC × investment = 10" are always insignificant.

The estimations reported in Table S3 and Table S4 indicate a significant differential effect of rTMS across stimulations (right DLPFC, left DLPFC, sham) in the reputation condition, but not in the anonymous condition. Does the strength of this differential effect across stimulations also differ significantly *between* the 2 experimental conditions, reputation and anonymous? To investigate this question, we pooled *all* data from both experimental conditions (reputation and anonymous). We used the same 4 regression specifications as in Table S3 and Table S4, and also interacted all variables with a dummy variable, "reputa-

- Robertson EM, Théoret H, Pascual-Leone A (2003) Studies in cognition: The problems solved and created by transcranial magnetic stimulation. J Cogn Neurosci 15:948–960.
- Wout MV, Kahn R, Sanfey AG, Aleman A (2005) Repetitive transcranial magnetic stimulation over the right dorsolateral prefrontal cortex affects strategic decisionmaking. NeuroReport 16:1849–1852.
- Koch G, et al. (2005) rTMS evidence of different delay and decision processes in a frontoparietal neuronal network activated during spatial working memory. Neuroimage 24:34–39.
- Griskova I, Ruksenas O, Dapsys K, Herpertz S, Hoppner J (2007) The effects of 10-Hz transcranial magnetic stimulation on resting EEG power spectrum in healthy subjects. Neurosci Lett 419:162–167.

tion," that indicates the experimental condition (i.e., whether or not an observation stems from the reputation condition). We report the corresponding estimation results in Table S5. We see that the interaction of the factors "reputation \times right DLPFC \times investment = 10" is significant and negative in all specifications, regardless of which control variables we include (P = .01 fordifferences in right DLPFC vs. sham stimulation effects between experimental conditions in both specifications 1 and 3; P < .01for differences in right DLPFC vs. left DLPFC stimulation effects between experimental conditions in both specifications 2 and 4). In contrast, the interactions for the other stimulations "reputation \times left DLPFC \times investment = 10" and "reputation × sham × investment = 10" are insignificant in all specifications. Finally, note that the coefficient of the dummy variable "reputation" is always positive (0.25 in both specifications 1 and 2, and 0.22 in both specifications 3 and 4) and always significant. This indicates that average back-transfers in the reputation condition were 22–25 percentage points higher in the reputation condition than in the anonymous condition; that is, trustees cared greatly about their reputation when reputation formation was possible.

- 5. Wasserman EM (1998) Risk and safety of repetitive transcranial magnetic stimulation. Electroencephalogr Clin Neurophysiol 108:1–16.
- Carver CS, White TL (1994) Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS scales. J Pers Soc Psychol 67:319–333.
- Strobel A, Beauducel A, Debener S, Brocke B (2001) Eine deutschprachige Version des BIS/BAS-Fragebogens von Carver und White. Zeitschrift Differ Diagn Psychol 22:216– 227
- 8. Perugini M, Callucci M, Presaghi F, Ercolani AP (2003) The personal norm of reciprocity. Eur J Pers 17:251–283.



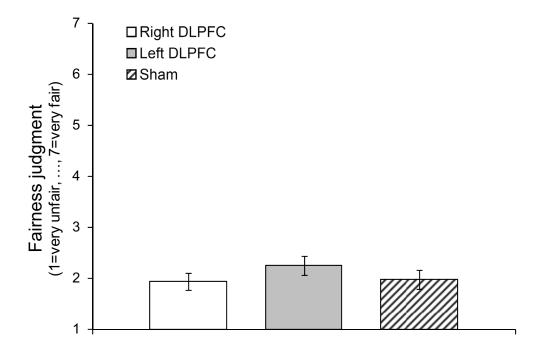


Fig. S1. Trustees' fairness ratings. Subjects assessed the fairness of a hypothetical trustee who returns nothing in response to an investor transfer of 7 points on a 7-point Likert scale. Means (all P > .16, Mann-Whitney tests) and distributions (P = .38, Kruskal-Wallis test) do not differ significantly across the stimulation groups.

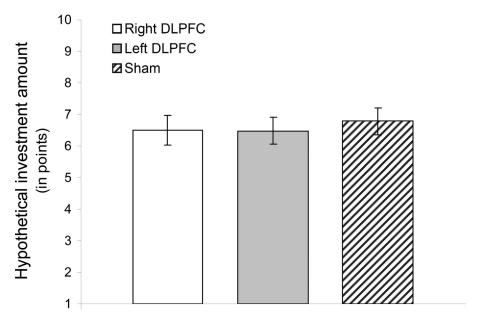


Fig. S2. Trustees' assessment of the future consequences of a profile of past back-transfers. Subjects predicted the likely transfer of a hypothetical investor who observes that his current trustee opted twice for "equalize payoff" and once for back-transferring nothing in the previous 3 periods. Means (all P > .78, Mann-Whitney tests) and distributions (P = .95, Kruskal-Wallis tests) do not differ significantly across the stimulation groups.

Table S1. Number of participants by treatment

	Right DLPFC	Left DLPFC	Sham	Total
Anonymous	15	14	14	43
Reputation	15	15	14	44
Total	30	29	28	87

Table S2. Payoff structure

Trustee's choice

Investment	Transfer nothing		Transfer 25% of received amount		Equalize payoffs	
	Investor	Trustee	Investor	Trustee	Investor	Trustee
1 point	9	14	10	13	11.5	11.5
4 points	6	26	10	22	16	16
7 points	3	38	10	31	20.5	20.5
10 points	0	50	10	40	25	25

Table S3. Reputation condition: GLS regression of the back-transfer decisions on indicators of the experimental conditions and other controls

Reputation condition 2 3 1 4 Right DLPFC -0.00295-0.00295-0.00789-0.00789(0.03479)(0.03479)(0.03565)(0.03565)Left DLPFC -0.02785-0.027850.00413 0.00413 (0.03697)(0.03697)(0.03714)(0.03714)Investment = 10 -0.10216^{\dagger} -0.01789 -0.10732^{\dagger} -0.03952(0.03321)(0.03352)(0.03334)(0.03305)Right DLPFC \times Investment = 10 -0.11380^{\dagger} -0.19807^{\dagger} -0.11095^{\dagger} -0.17875^{\dagger} (0.04669)(0.04692)(0.04646)(0.04644)Left DLPFC \times Investment = 10 0.08427 0.06780 (0.04719)(0.04674) $Sham \times Investment = 10$ -0.08427-0.06780(0.04719)(0.04674)Fairness judgment -0.06128^{\dagger} -0.06128^{\dagger} (0.01637)(0.01637)-0.00864Impulsivity (BIS) -0.00864(0.04164)(0.04164)Impulsivity (BAS) 0.00629 0.00629 (0.04383)(0.04383)Reciprocity (positive) 0.04148 0.04148 (0.02186)(0.02186)Reciprocity (negative) 0.01750 0.01750 (0.01846)(0.01846)Constant 0.49751^{\dagger} 0.49751^{\dagger} 0.35583* 0.35583* (0.02550)(0.02550)(0.16407)(0.16407)Observations 616/44 616/44 616/44 616/44 0.13 R2 (within) 0.13 0.13 0.13

Dependent variable: fraction of received points that the trustee transfers back. Robust SEs, adjusted for clustering on subject level, are given in parentheses. *Significant at 5%.

[†]Significant at 2%.

Table S4. Anonymous condition: GLS regression of the back-transfer decisions on indicators of the experimental conditions and other controls

Anonymous condition 1 2 3 4 Right DLPFC -0.04169-0.04169-0.08184-0.08184(0.05937)(0.05937)(0.06275)(0.06275)Left DLPFC -0.02250-0.02250-0.03246-0.03246(0.06082)(0.06082)(0.06403)(0.06403)Investment = 10 0.00446 0.05210 0.00564 0.054225 (0.03857)(0.03543)(0.03879)(0.04714)Right DLPFC \times Investment = 10 0.05848 0.01085 0.06325 0.014666 (0.04898)(0.04655)(0.04980)(0.04714)Left DLPFC \times Investment = 10 0.04764 0.04859 (0.05237)(0.05237) $Sham \times Investment = 10$ -0.04764-0.04859(0.05237)(0.05237)Fairness judgment -0.03649^{\dagger} -0.03649^{\dagger} (0.01884)(0.01884)Impulsivity (BIS) 0.02246 0.02246 (0.07246)(0.07246)Impulsivity (BAS) -0.07697-0.07697(0.06358)(0.06358)Reciprocity (positive) 0.00427 0.00427 (0.04176)(0.04176)Reciprocity (negative) -0.00770-0.00770(0.04579)(0.04579)Constant 0.25549^{\dagger} 0.25549^{\dagger} 0.50856 0.50856 (0.04576)(0.04576)(0.35623)(0.35623)602/43 602/43 Observations/groups 588/42 588/42 R2 (within) 0.01 0.01 0.01 0.01

Dependent variable: fraction of received points that the trustee transfers back. Robust SEs are in parentheses, adjusted for clustering on subject level. †Significant at 2%.

Table S5. Pooled estimation: Reputation condition and anonymous condition. GLS regression of the back-transfer decisions on indicators of the experimental conditions and other controls

Pooled model 1 2 3 4 Right DLPFC -0.04243-0.04243-0.07975-0.07975(0.05028)(0.05028)(0.05026)(0.05026)Left DLPFC -0.02350-0.02350-0.03276-0.03276(0.05126)(0.05130)(0.05130)(0.05126)Investment = 10 0.00156 0.00156 0.00259 0.00259 (0.03884)(0.03884)(0.03893)(0.03893) $Right\ DLPFC \times Investment = 10$ 0.06078 0.06078 0.06490 0.06490 (0.04971)(0.04971)(0.05051)(0.05051) $Left\ DLPFC \times Investment = 10$ 0.05055 0.05055 0.05377 0.05377 (0.05291)(0.05291)(0.05269)(0.05269)Reputation 0.25154 0.25154 0.22103[†] 0.22103^{\dagger} (0.05115)(0.05115)(0.05028)(0.05028)Reputation \times Right DLPFC 0.03685 0.03685 0.07298 0.07298 (0.06810)(0.06810)(0.06753)(0.06753)Reputation \times Left DLPFC 0.00145 0.00145 0.03424 0.03424 (0.07000)(0.07000)(0.06919)(0.06919)Reputation \times Investment = 10 -0.12475^{\dagger} -0.10497* -0.12663^{\dagger} -0.11446^{\dagger} (0.05101)(0.04877)(0.05100)(0.04848)Reputation \times Right DLPFC \times Investment = 10 -0.17306^{\dagger} -0.19284^{\dagger} -0.17430^{\dagger} -0.18647^{\dagger} (0.06753)(0.06585)(0.06804)(0.06617)Reputation \times Left DLPFC \times Investment = 10 0.01978 0.01217 (0.07057)(0.07029)Reputation \times Sham \times Investment = 10 -0.01978-0.01217(0.07057)(0.07029)Fairness judgment -0.04238^{\dagger} -0.04238^{\dagger} (0.01227)(0.01227)Impulsivity (BIS) -0.00041-0.00041(0.04031)(0.04031)Impulsivity (BAS) -0.04067-0.04067(0.03903)(0.03903)Reciprocity (positive) 0.02919 0.02919 (0.02196)(0.02196)0.01408 Reciprocity (negative) 0.01408 (0.01993)(0.01993)Constant 0.25648 0.25648 0.27945 0.27945 (0.03852)(0.03852)(0.17066)(0.17066)Observations/groups 1,218/87 1,218/87 1,204/86 1,204/86 R2 (within) 0.07 0.07 0.08 0.08

Dependent variable: fraction of received points that the trustee transfers back. Robust SEs, adjusted for clustering on subject level, are given in parentheses. *Significant at 5%.

[†]Significant at 2%